Cybernetics and Systems Analysis, Vol. 36, No. 3, 2000

SYSTEMS ANALYSIS

CHARACTERISTIC OF EXIT MOMENTS AND MODELS OF ENLARGEMENT OF STATES FOR FINITE MARKOV CHAINS IN TERMS OF GLOBAL MEMORY FUNCTIONALS

V. V. Anisimov[a] **and Yu. Yu. Klygunova**[b]

UDC 519.21

A notion of local- and global-memory functionals for discrete-type distributions is introduced by analogy with the notion of the memory for continuous-type distributions introduced in Muth's papers. In the class of PH-distributions (i.e., of distributions of the time of Markov chain exit from a subset of states) the necessary and sufficient conditions are obtained for the case where the exit time has an exponential (continuous-time) or a geometric (discrete time) distribution. A new notion of a global memory functional for decomposition of the state space of a finite Markov chain is introduced. Its properties as a measure of quality of decomposition and enlargement of a state space are studied. The asymptotic optimality is proved.

Keywords: *exit moment, model of enlargement of states, finite Markov chain, global memory functional.*

1. CHARACTERIZATION OF *PH*-DISTRIBUTIONS IN TERMS OF MEMORY FUNCTIONALS

1. Memory Functionals for Discrete Distributions. It is well known that the probability characteristics of any random variable are determined by its distribution function or (in the absolutely continuous case) by its density function. For the analysis of non-negative random variables, Barlow and Proshan have applied the function of failure rates. In terms of this function, it is more convenient to characterize such effects as aging. The function of average remainder of lifetime is an integral analog of the failure function and depends on operation time. It is determined as

$$r(t)=\frac{1}{R(t)}\int_t^{\infty} R(x)dx, \tag{1}$$

where $R(t)=P(T>t)$.

Based on of this function, the American scientist Muth [7] has introduced a new characteristic of probability distributions — the memory. The term "memory" appeared due to the fact that the absence of aftereffect is characteristic of exponential distribution, i.e., this distribution has zero memory. Therefore, the memory characteristic determines in some sense the degree of presence of an aftereffect. Muth has constructed the functionals of local and global memory for continuous distributions

---

[a]Bilkent University, Ankara, Turkey and Taras Shevchenko University, Kiev, Ukraine. [b]Wane University, Michigan, USA. Translated from Kibernetika i Sistemnyi Analiz, No. 3, pp. 110-120, May-June, 2000. Original article submitted September 18, 1998.

1060-0396/00/3603-0405$25.00 ©2000 Kluwer Academic/Plenum Publishers

$$m_L = -\frac{dr(t)}{dt}, \quad m_G = 2 - 2\int_0^\infty r(x)R(x)dx,$$

or

$$m_G = 2 - \frac{ET^2}{(ET)^2}, \tag{2}$$

He also has classified them by the values of m, according to the following four types:

1) if $m = 1$, then the distribution has a perfect memory;

2) if $m = 0$, then the distribution has a zero memory;

3) if $m \in (-\infty, 0)$, then the distribution has a negative memory;

4) if $m \in (0, 1)$, then the distribution has a positive memory.

Let us construct functionals of local and global memory for discrete distributions. We will consider a random variable T, which takes on values from the set $\{1, 2, \dots\}$. Let $p_k = P(T = k)$. By analogy with (1), let us define the function of average remainder of life:

$$r(t) = \frac{\sum_{l \geq k} P(T > l)}{P(T > k)}, \quad k = 0, 1, \dots \tag{3}$$

Note that $r(0) = \mu = \mathbf{E}T$ since

$$r(0) = \sum_{l \geq 0} \sum_{i > l} p_i = \sum_{i \geq 0} \sum_{l \geq 0}^{i-1} p_i = \sum_{i \geq 0} ip_i = \mu. \tag{4}$$

Let us define local memory at a point.

Definition 1. By the local memory of a random variable T assuming values $\{1, 2, \dots\}$, we will mean the quantity

$$m_L(k) = r(k) - r(k+1), \quad k \geq 0.$$

Let us define the functional of global memory for the discrete quantity T. The general form of the global memory functional for T is determined as

$$m_G = \sum_{k \geq 0} m_L(k) w(k), \tag{5}$$

where $w(k)$ is some weight function. Let us transform expression (5) designating $\tilde{r}(k) = \mu - r(k)$

$$\sum_{k \geq 0} (r(k) - r(k+1)) w(k) = -\sum_{k \geq 0} \tilde{r}(k) w(k+1) + \sum_{k \geq 1} \tilde{r}(k) w(k)$$

$$= \sum_{k \geq 1} \tilde{r}(k)(w(k) - w(k+1)) - \tilde{r}(0) w(1). \tag{6}$$

We will select the weight function $w(k)$ so that the geometrical distribution has zero global memory, i.e., the property of absence of aftereffect in the discrete analog of an exponential distribution is fulfilled. This condition is satisfied by the function

$$w(k) = \frac{2}{\mu^2} \sum_{j \geq k} P(T > j).$$

Since $\tilde{r}(0) = 0$, and

$$w(k) - w(k+1) = \frac{2}{\mu^2} P(T > k),$$

we will rewrite (6) in the form

$$m_G = \sum_{k \geq 0} (\mu - r(k)) \frac{2}{\mu^2} P(T > k) = \frac{2}{\mu} \sum_{k \geq 0} P(T > k) - \frac{2}{\mu^2} \sum_{k \geq 0} \sum_{l \geq k} P(T > l). \tag{7}$$

The first term on the right-hand side of (7) is equal to two according to (4). For the second term, we will obtain

$$\sum_{k\ge 0}\sum_{l\ge k} P(T>l)=\sum_{k\ge 0}\sum_{l\ge k}\sum_{i\ge l+1} p_i=\sum_{k\ge 0}\sum_{i\ge k+1}\sum_{l=k}^{i-1} p_i$$

$$=\sum_{i\ge 1}\left(i^2 p_i - p_i\sum_{k=0}^{i-1} k\right)=\sum_{i\ge 1}\left(i^2-\frac{i(i+1)}{2}\right)p_i=\frac{\mathbf{E}T^2+\mathbf{E}T}{2}.$$

Finally, we will obtain the following definition for global memory.

Definition 2. By the global memory of a random variable T assuming values from the set $\{1,2,\dots\}$, we will mean

$$m_G=2-\frac{\mathbf{E}T^2+\mathbf{E}T}{(\mathbf{E}T)^2}. \tag{8}$$

Checking the latter expression for the geometric distribution $p_i=P(T=i)=pq^{i-1}$ for all $i\ge 1$, $p\in(0,1)$, we will obtain that the global memory is equal to zero.

It should be pointed out that if the distribution is geometric or exponential, then the global memory is zero, but the converse is generally incorrect.

2. Characterization of *PH*-Distributions in Terms of Global Memory Functionals. Since global memory functionals characterize distributions in terms of the first two moments, i.e., it is a numerical and a rather simply calculated characteristic, of interest is a search for classes of distribution characterized in terms of this functional. An answer to this problem is obtained to some extent for so-called *PH*-distributions (distributions characterizing the time of stay of a Markovian process in a subset of states).

Let us consider a Markov chain with the discrete time ξ_m, $m\ge 0$, with the finite set of states X, $\dim X=n$, and with the initial distribution vector $\bar p=(p_m)_{m=\overline{1,n}}$ and the matrix of transition probabilities $P=\|p_{i,j}\|_{i,j=\overline{1,n}}$.

Let τ_i be the time of stay of the chain ξ_m in the subset $I\subset X$, $\dim I=d<n$, up to the moment of the first exit provided that the initial state $i\in I$. We will write the system of stochastic relations for τ_i

$$\tau_i=1+\sum_{j\in I} I^j_{x_1}\tau_j,$$

where

$$I^j_k=\begin{cases}1, & j=k,\\ 0, & j\ne k.\end{cases}$$

Passing in the latter equality to the mathematical expectation and to the second moment, we will obtain the relations

$$m_i=1+\sum_{j\in I} p_{ij}m_j, \tag{9}$$

$$d_i=1+2\sum_{j\in I} p_{ij}m_j+\sum_{j\in I} p_{ij}d_j, \tag{10}$$

where $m_i=\mathbf{E}\tau_i$, $d_i=\mathbf{E}\tau_i^2$.

Let us introduce a substochastic matrix on the subset I

$$\tilde P=\|p_{i,j}\|_{i,j\in I}.$$

The following theorem characterizes *PH*-distributions in terms of global memory functional.

THEOREM 1. Let I be a nonreducible and nonclosed subset. For the distribution of the time of stay τ_i in the subset I not to depend on the initial distribution and be geometrical, it is necessary and sufficient that for any $i\in I$: $m_G(\tau_i)=0$, i.e.,

$$\mathbf{E}\tau_i^2+\mathbf{E}\tau_i-2(\mathbf{E}\tau_i)^2=0. \tag{11}$$

Proof. The necessity is obvious.

Sufficiency. Let conditions (11) be satisfied for any $i\in I$, where I is a nonclosed and nonreducible subset on which a substochastic matrix $\tilde P$ is defined. Then such matrix is nonreducible and there is a matrix $(I-\tilde P)^{-1}$.

Let us introduce the following notation:

$$\overline{m} = \begin{bmatrix} m_1 \\ m_2 \\ \cdot \\ \cdot \\ \cdot \\ m_k \end{bmatrix}, \quad \overline{d} = \begin{bmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ \cdot \\ d_k \end{bmatrix}, \quad 1 = \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}, \quad I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

and let us write Eqs. (9), (10) in the vector form

$$\overline{m} = 1 + \widetilde{P}\,\overline{m}, \tag{12}$$

$$\overline{d} = (I - \widetilde{P})^{-1}(1 + 2\widetilde{P}(I - \widetilde{P})^{-1}1). \tag{13}$$

Let us determine the diagonal matrix A

$$A = \begin{pmatrix} m_1 & 0 & \dots & 0 \\ 0 & m_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & m_k \end{pmatrix}$$

and let us write in terms of $A\overline{m}$ the column vector $A\overline{m} = ((\mathbf{E}\tau_i)^2,\ i \in I)$. Then as a result of transformations, Eq. (11) with allowance for (12), (13) will have the form

$$(I + (I - \widetilde{P})^{-1}\widetilde{P} - A)(I - \widetilde{P})^{-1}1 = 0.$$

After its multiplication by the matrix $I - \widetilde{P}$, we will obtain

$$(I + \widetilde{P}A - A)(I - \widetilde{P})^{-1}1 = 0. \tag{14}$$

Let us prove that if the matrix $\widetilde{P}$ is nonreducible, then from Eqs. (12), (14) it follows that

$$\overline{m} = (I - \widetilde{P})^{-1}1 = c1,$$

where $c = \text{const} > 0$. If it is so, then

$$c(I - \widetilde{P})1 = 1,$$

whence $\overline{q} = c^{-1}1$, where $q_i = \sum_{j \notin I} p_{ij}$, i.e., for all $i \in I$

$$\sum_{j \notin I} p_{ij} = \text{const}.$$

Thus, the time of stay τ_i is geometrically distributed.

Denote by $(I - \widetilde{P})^{-1}1 = \overline{c}$, where $\overline{c} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_k \end{bmatrix}$.

We will write the system of equations for the quantities c_i. From Eq. (14) it follows that

$$\bar{c} - (I - \tilde{P})A\bar{c} = 0,$$

i.e.,

$$c_i - c_i m_i + \sum_{j \in I} p_{ij} m_j c_j = 0. \tag{15}$$

Denote $D = I - \tilde{P}$ and $m_i c_i = \tilde{m}_i$ for $i \in I$. Let $\overline{m}$ be a column vector with the elements $\tilde{m}_i$, $i \in I$. Then systems (11), (15) will have the form

$$\begin{cases} D\overline{m} = 1, \\ D\tilde{m} = \bar{c}, \end{cases}$$

whence

$$\begin{cases} \overline{m} = D^{-1}1, \\ \tilde{m} = D^{-1}\bar{c}. \end{cases} \tag{16}$$

Let

$$D^{-1} = (I - \tilde{P})^{-1} = \|\alpha_{i,j}\|_{i,j \in I};$$

then

$$D^{-1}1 = \left(\sum_{j \in I} \alpha_{ij},\ i \in I \right).$$

It follows from system (16) that

$$\begin{cases} m_i = \sum_{j \in I} \alpha_{ij}, \\ m_i c_i = \sum_{j \in I} \alpha_{ij} c_j \end{cases}$$

or

$$c_i = \frac{\sum_{j \in I} \alpha_{ij} c_j}{\sum_{j \in I} \alpha_{ij}}, \quad i \in I. \tag{17}$$

Let us determine the matrix G:

$$G = \left\| \frac{\alpha_{ij}}{\sum_j \alpha_{ij}} \right\|.$$

Then we can rewrite system (17) as follows:

$$\bar{c} = G\bar{c}. \tag{18}$$

Let us study the properties of the matrix G. First let us note that the matrix D^{-1} has positive elements. Indeed, $\tilde{P}$ is a substochastic matrix and the absolute value of all its eigenvalues is strictly less than unity. Then according to [3, p. 367], all elements of the matrix $(I - \tilde{P})^{-1}$ are positive. Let us prove that if the matrix $\tilde{P}$ is nonreducible, then G is also nonreducible. To do this, it will suffice to prove irreducibility of $(I - \tilde{P})^{-1}$. Let $\tilde{P}$ be nonreducible and $(I - \tilde{P})^{-1}$ be reducible, i.e., without loss of generality it has the form

$$(I - \tilde{P})^{-1} = \begin{pmatrix} A_1 & 0 \\ C & A_2 \end{pmatrix}.$$

Then according to [3, p. 61],

$$(I-\tilde{P})=\begin{pmatrix} A_1^{-1} & 0 \\ -A_2^{-1}CA_1^{-1} & A_2^{-1} \end{pmatrix},$$

whence

$$\tilde{P}=\begin{pmatrix} I-A_1^{-1} & 0 \\ -A_2^{-1}CA_1^{-1} & I-A_2^{-1} \end{pmatrix}. \qquad \square$$

Thus, $\tilde{P}$ is reducible and this contradicts the condition of the theorem, where $\tilde{P}$ is nonreducible. Therefore, $(I-\tilde{P})^{-1}$ is nonreducible, and because of this G is also nonreducible. Since all the elements of G are positive and $\sum \alpha_{ij} > 0$, by the construction, G is a nonreducible stochastic matrix. Therefore, according to [3], it has a unique eigenvalue equal to unity, to which eigenvector 1 corresponds. In other words, the equation $G\bar{x}=\bar{x}$ has a unique solution to within a constant. It follows herefrom that the vectors $\bar{c}$ have the form $\bar{c}=c1$ in system (18). This is what we had to prove.

Let us consider a regular Markovian process with the continuous time $\xi(t)$, $t \geq 0$, in the phase space X, $\dim X = n$, which is given

1) by the matrix of transition probabilities of the embedded Markov chain

$$P=\|p_{ij}\|_{i,\,j=\overline{1,n}}=\|\lambda_{ij}\,\lambda_i^{-1}\|_{i,\,j=\overline{1,n}};$$

2) by the vector $\{\theta_i, i=\overline{1,n}\}$ of times of stay in the states having exponential distribution:

$$\mathbf{P}\{\theta_i < x\}=1-e^{-\lambda_i x},\ i=\overline{1,n}\ ;$$

3) by the vector of the initial distribution

$$\bar{p}=\{p_i(0),\ i=\overline{1,n}\}.$$

Let $I \subset X$; $\dim I = d < n$; τ_i be the time of stay of $\xi(t)$ in the subset I up to the moment of the first exit provided that the initial state $i \in I$. Denote by ξ_k, $k \geq 0$, the embedded Markov chain for $\xi(t)$. Then the values of τ_i satisfy the system of stochastic relations

$$\tau_i=\theta_i+\sum_{j\in I} I_{ij}\tau_j, \tag{19}$$

where I_{ij} is the indicator of passage from i to j for the embedded chain ξ_k at the first step provided that $\xi_0 = i \in I$.

Passing in (19) to the mathematical expectation and to the second moment, we will obtain the following systems of equations:

$$m_i=\frac{1}{\lambda_i}+\sum_{j\in I} p_{ij}m_j,$$

$$d_i=\frac{2}{\lambda_i^2}+2\sum_{j\in I} p_{ij}m_j\frac{1}{\lambda_j}+\sum_{j\in I} p_{ij}d_j,$$

where $m_i=\mathbf{E}\tau_i$, $d_i=\mathbf{E}\tau_i^2$. Denoting by $\tilde{P}=(p_{ij})_{i,\,j\in I}$ a substochastic matrix, which is defined on the subset I, and introducing the quantities

$$\bar{\lambda}=\begin{Bmatrix} 1/\lambda_1 \\ 1/\lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ 1/\lambda_k \end{Bmatrix},\quad \Lambda=\begin{pmatrix} 1/\lambda_1 & 0 & \dots & 0 \\ 0 & 1/\lambda_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1/\lambda_k \end{pmatrix},$$

we will obtain the vector equations

$$\overline{m} = \overline{\lambda} + \tilde{P}\,\overline{m}, \tag{20}$$

$$\overline{d} = 2\Lambda\overline{\lambda} + 2\Lambda\tilde{P}\,\overline{m} + \tilde{P}\,\overline{d}. \tag{21}$$

Then the following theorem will be true.

THEOREM 2. Let the substochastic matrix $\tilde{P}$ be defined on the subset I, which is nonclosed and nonreducible (i.e., it does not contain closed subsets). Then for the distribution of the time of stay τ_i in the subset I not to depend on the initial state and be exponentially distributed, it is necessary and sufficient that for any $i \in I$: $m_G(\tau_i) = 0$, i.e.,

$$\mathbf{E}\tau_i^2 - 2(\mathbf{E}\tau_i)^2 = 0. \tag{22}$$

Proof. The structure of the proof is similar to the proof of Theorem 1. Having transformed (22) according to Eqs. (20), (21), we obtain the vector representation of condition (22)

$$(\Lambda - (I - \tilde{P})A)(I - \tilde{P})^{-1}\,\overline{\lambda} = 0, \tag{23}$$

where

$$A = \begin{pmatrix} m_1 & 0 & \dots & 0 \\ 0 & m_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & m_k \end{pmatrix}.$$

Further, by analogy with the proof of Theorem 1, we come to the conclusion that if $\tilde{P}$ is nonreducible, then from Eqs. (20), (23) it follows that

$$\overline{m} = (I - \tilde{P})^{-1}\,\overline{\lambda} = c\,1,$$

where $c = \text{const} > 0$. If this is so, then

$$2c(I - \tilde{P})\,1 = \overline{\lambda}, \quad \overline{q} = c^{-1}\,\overline{\lambda},$$

where $q_i = \sum_{j \notin I} p_{ij}$, i.e., for all $i \in I$

$$\sum_{j \notin I} \lambda_{ij} = \text{const} = \frac{1}{c}.$$

Therefore, the time of stay τ_i in the subset will be exponential with the parameter c^{-1}, since the intensities of exit from the subset for each state will be identical and equal to c^{-1}. □

2. GLOBAL MEMORY FUNCTIONAL FOR PARTITION AND ENLARGEMENT OF A PHASE SPACE OF MARKOVIAN PROCESSES

1. Global Memory Functional for Partition. The results of Theorems 1 and 2 are the theoretical foundation for a new approach to the problems of the analysis of partition and enlargement of states of Markovian processes. Let us construct a global memory functional for partition of phase space of the Markovian process.

Denote by Γ the set of all partitions γ dividing a finite set of states of phase space of a Markovian process X, $\dim X = n$ into non-intersecting subsets and whose join composes the whole set X, i.e., $\gamma = \{I_\alpha, 1 \le \alpha \le n_\gamma\}$, where $1 \le n_\gamma \le n$. Let τ_i^α be the time of stay of a Markovian process in the subset I_α provided that the initial state $i \in I_\alpha$.

Definition 3. By the global memory functional for the partition γ we mean the expression

$$F(\gamma) = \sum_{\alpha=1}^{n_\gamma} \max_{i \in I_\alpha} |m_G(\tau_i^\alpha)|.$$

The functional constructed can be considered a measure of the partition quality, namely: the less the value of the functional in the partition, the better is this partition from the point of view of enlargement of the process. A desired optimal partition can be obtained by minimizing this functional.

Definition 4. The partition γ_0 is optimal, if $\gamma_0 = \arg\min_{\gamma} F(\gamma)$.

Let us study the properties of the functional F. We will consider some partition $\gamma = \{I_\alpha, 1 \le \alpha \le n_\gamma\}$ such that each subset $I_\alpha \in \gamma$ is nonclosed, there is some subset $I_\beta \in \gamma$ which is reducible, and the remaining subsets I_α are nonreducible. The subset I_β is reducible in the sense that $I_\beta = \bigcup_{\beta'} I_{\beta'}$, where $I_{\beta'}$ are noninterconnected subsets. Then the lemma on refinement will be true.

LEMMA 1 (On refinement). Let the following conditions be fulfilled:

1) in the partition $\gamma = \{I_\alpha, 1 \le \alpha \le n_\gamma\}$, any subset of the set of states I_α is nonclosed and there is a subset I_β reducible in the above-mentioned sense. The remaining I_α, $\alpha \ne \beta$, are nonreducible;

2) $F(\gamma) = 0$.

Then the partition γ can be refined to the partition γ' so that $F(\gamma') = 0$.

The partition γ' is constructed in such a manner that the reducible subset I_β is split into nonreducible subsets $\{I_\beta'\}$ so that $I_\beta = \bigcup_{\beta'} I_{\beta'}$, and the remaining subsets do not vary.

Remark 1. Note that the lemma will also be true if states exist in the subset I_β, which are inessential with respect to the remaining states from I_β.

Proof. If $F(\gamma) = 0$, then by the construction of F, $m_G(\tau_i^\alpha) = 0$ on each subset I_α from the partition γ. Without loss of generality, let us consider the reducible subset I_β, which can be reduced to nonreducible noninterconnecting subsets $I_{\beta'}$, $I_{\beta''}$, where $I_\beta = I_{\beta'} \bigcup I_{\beta''}$. Then the probabilities or intensities of exit from $I_{\beta'}$ and $I_{\beta''}$ will be identical constants. Therefore $m_G(\tau_i^{\beta'}) = 0$, $m_G(\tau_i^{\beta''}) = 0$.

Then if we designate the partition $\gamma' = (I_{\beta'}, I_{\beta''}, I_{\alpha_1}, \dots, I_{\alpha_k})$ by γ', then it is obvious that $F(\gamma') = 0$. □

Let us formulate the sufficient conditions of enlargement in terms of global memory functionals for a partition. Note that if the distributions of times of stay in partition subsets do not depend on the initial states and are geometrically or exponentially distributed, then it is obvious that $F(\gamma) = 0$.

LEMMA 2. Let on some partition γ such that the subset $I_\alpha \in \gamma$ is nonclosed and nonreducible, the condition $F(\gamma) = 0$ be satisfied. Then either the quantities $\sum_{j \notin I_\alpha} p_{ij}$ (in the case of discrete time) or the quantities $\sum_{j \notin I_\alpha} \lambda_{ij}$ (in the case of continuous time) do not depend on the initial state $i \in I_\alpha \in \gamma$, i.e., the values of τ_i^α have either geometrical or exponential distribution.

Proof. If all the nonclosed subsets I_α are nonreducible, then it follows from the construction of F and sufficient conditions of Theorems 1, 2 that the distribution of τ_i^α does not depend on the initial state and is geometrical or exponential. □

It should be noted, that the Kemeni–Snell enlargement conditions are stronger as compared with enlargement conditions in terms of global memory functional for a partition.

LEMMA 3. If the Markovian process satisfies on the partition γ the Kemeni–Snell enlargement conditions, then $F(\gamma) = 0$.

2. Asymptotic Properties of the Global Memory Functional for a Partition. Let us consider a perturbed Markovian process $\xi_\varepsilon(t)$, $t \ge 0$, depending on a small parameter ε and satisfying the asymptotic enlargement conditions, i.e., the matrix of transition probabilities of the embedded Markovian chain can be presented as $P_\varepsilon = P_0 + \varepsilon P_1$, where P_0 is a block diagonal matrix, and P_1 is the perturbation matrix. According to [1, 5, 6], a partition of phase space X such that

$$X = \bigcup_\alpha I_\alpha, \quad I_\alpha \cap I_\beta = \varnothing, \ \alpha \ne \beta, \tag{24}$$

corresponds to such representation of the matrix P_ε. Here, for the matrix P_0 , we have $p_0(i, j) = 0$, $i \in I_\alpha$, $j \in I_\beta$, and the classes I_α for a nonperturbed Markov process are essential.

Denote by γ_0 a partition satisfying the conditions of asymptotic enlargement (24). Let $\tau_{i,\alpha}^\varepsilon$ be the time of stay $\xi_\varepsilon(t)$ of such partition in the subset I_α up to the moment of the first exit provided that $\xi_\varepsilon(0) = i \in I_\alpha$.

The following theorem describes the asymptotic properties of a global memory functional up to the moment of exit from one class of such partition γ_0.

THEOREM 3. If for the perturbed process $\xi_\varepsilon(t)$ there exists a partition satisfying asymptotic integrability conditions, then $m_G(\tau^\varepsilon_{i,\alpha}) = O(\varepsilon) \to 0$ for $\varepsilon \to 0$ on each class of such partition and the time of stay in each class has (with the respective normalization) an asymptotically exponential distribution.

Proof. Let us consider one class $I_\alpha \in \gamma_0$ with the matrix of transition probabilities $\tilde{P}_\varepsilon = \| p^\varepsilon_{ij} \|_{i,\, j \in I_\alpha}$. Since $\tilde{P}_\varepsilon \to \tilde{P}_0$ for $\varepsilon \to 0$, where $\tilde{P}_0$ is a stochastic and nonreducible matrix, according to [2], the representation

$$(I - \tilde{P}_\varepsilon)^{-1} = \frac{1}{\varepsilon\beta}\Pi + C_\varepsilon$$

holds, where Π is a matrix with identical rows of the form $(\pi_1, \ldots, \pi_k)$; π_i are stationary probabilities for the matrix $\tilde{P}_0$, and

$$\beta = \sum_{j \in I_\alpha} \pi_j \beta_j,$$

$$\beta_j = \lim_{\varepsilon \to 0} \varepsilon^{-1} \sum_{k \notin I_\alpha} p_\varepsilon(j,k), \quad j \in I_\alpha,$$

and the elements of the matrix $C_\varepsilon = \| C^\varepsilon_{ij} \|_{i,\, j \in I_\alpha}$ have finite limits for $\varepsilon \to 0$.

Let $\xi_\varepsilon(0) = i \in I_\alpha$. Then according to (20), (21),

$$m_G(\tau^\varepsilon_{i,\alpha}) = 2\left[1 - \left(\frac{1}{\varepsilon\beta}\sum_j \pi_j \lambda_j^{-1}\left(\frac{1}{\varepsilon\beta}\sum_k \pi_k \lambda_k^{-1} + \sum_k C^\varepsilon_{kj}\lambda_k^{-1}\right)\right.\right.$$

$$\left.\left. + \sum_j C^\varepsilon_{ij}\lambda_j^{-1}\left(\frac{1}{\varepsilon\beta}\sum_k \pi_k\lambda_k^{-1} + \sum_k C^\varepsilon_{kj}\lambda_k^{-1}\right)\right)\left(\frac{1}{\varepsilon\beta}\sum_j \pi_j\lambda_j^{-1} + \sum_j C^\varepsilon_{ij}\lambda_j^{-1}\right)^{-2}\right].$$

Denote by $T = \sum_{j \in I_\alpha} \pi_j \lambda_j^{-1}$ the stationary mathematical expectation for the process in the subset I_α. Then it follows from the latter formula that

$$m_G(\tau^\varepsilon_{i,\alpha}) = 2\varepsilon\beta\left[2T\sum_j C^\varepsilon_{ij}\lambda_j^{-1} + \varepsilon\beta\left(\sum_j C^\varepsilon_{ij}\lambda_j^{-1}\right)^2 - \sum_j \pi^\varepsilon_j \lambda_j^{-1}\sum_k C^\varepsilon_{kj}\lambda_k^{-1}\right.$$

$$\left. + \sum_j C^\varepsilon_{ij}\lambda_j^{-1}\left(T + \varepsilon\beta\sum_k C^\varepsilon_{kj}\lambda_k^{-1}\right)\right] \times \left(T + \varepsilon\beta\sum_j C^\varepsilon_{ij}\lambda_j^{-1}\right)^{-2}.$$

From here, $m_G(\tau^\varepsilon_{i,\alpha}) = 0(\varepsilon) \to 0$ for $\varepsilon \to 0$. □

It follows immediately from the theorem that the global memory functional for a partition on a partition satisfying the asymptotic enlargement conditions also tends to zero.

Denote by F_ε a global memory functional for partition of the perturbed process $\xi_\varepsilon(t)$. Then the following theorem, being a constructive algorithm for construction of an optimal partition, will be true.

THEOREM 4. If there is a unique partition γ_0 such that $\lim \lim_{\varepsilon \to 0} F_\varepsilon(\gamma_0) = 0$, and on any $\gamma \neq \gamma_0$: $\lim_{\varepsilon \to 0} F_\varepsilon(\gamma) > 0$, then there exists such ε_0 that for any $\varepsilon \le \varepsilon_0$ we have $\gamma_\varepsilon = \gamma_0$, where $\gamma_\varepsilon = \arg\min_{\beta \in \Gamma} F_\varepsilon(\beta)$ and $\gamma_0 = \arg\min_{\beta \in \Gamma} \lim_{\varepsilon \to 0} F_\varepsilon(\beta)$.

Proof. Since the set Γ is finite and for any $\gamma \neq \gamma_0$ $\lim_{\varepsilon \to 0} F_\varepsilon(\gamma) > 0$ and $\lim_{\varepsilon \to 0} F_\varepsilon(\gamma_0) = 0$, then there is a $\delta > 0$, such that, for $\gamma \neq \gamma_0$, beginning from some ε_0 for any $\varepsilon \le \varepsilon_0$, $F_\varepsilon(\gamma) > \delta$, which obviously proves our statement since $F_\varepsilon(\gamma_0) \to 0$. □

The obtained properties of a global memory functional for a partition allows us to consider it as a measure of partition quality, which is a new approach to the problems of construction of a partition optimal with respect to quality of partition of phase space enlargement.

REFERENCES

1. V. V. Anisimov, "Asymptotic enlargement of the states of random processes, " Kibernetika, No. 3, 109-117 (1973).
2. V. V. Anisimov, " Limiting theorems for the sums of random quantities on a Markovian chain associated with the exit from the set forming one class in a limit," Teor. Veroyatn. Mat. Stat., No. 4, 3-10 (1974).
3. F. R. Gantmakher, Theory of Matrices [in Russian], Nauka, Moscow (1967).
4. J. Kemeni and M. Snell, Finite Markov Chains [Russian translation], Nauka, Moscow (1970).
5. V. S. Korolyuk and A. F. Turbin, Mathematical Fundamentals of Phase Enlargement of Complex Systems [in Russian], Vyshcha Shk., Kiev (1978).
6. V. S. Korolyuk and A. F. Turbin, Semi-Markovian Processes and Their Applications [in Russian], Naukova Dumka, Kiev (1976).
7. E. J. Muth, "Memory as a property of probability distributions," IEEE. Trans. Reliabality, R-29, No. 2, pp. 160-165, (1980).